DOI:
10.1016/j.neubiorev.2018.12.028

Document Version
Peer reviewed version

Brain activation during human defensive behaviour: a systematic review and preliminary meta-analysis.

Fiona Patrick[1*], Matthew J Kempton[2,4], Lindsey Marwood[1], Steven C R Williams[3,4], Allan H Young[1,4], Adam Perkins[1,4]

[1] Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.

[2] Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.

[3] Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.

[4] National Institute for Health Research, Biomedical Research Centre, South London & Maudsley NHS Foundation Trust, London, UK.

*Corresponding author: Fiona Patrick (fiona.patrick@kcl.ac.uk)

Abstract

The neural underpinnings of defensive behaviour have implications for both basic research and clinical translation. This review systematically collates published research on neural response during simple avoidance of threat and approach-avoidance behaviour during goal-conflicting situations and presents an exploratory meta-analysis of available whole-brain data. Scopus, PsychInfo and Web of Science databases were searched for the period up to March 2018. 1,348 simple avoidance and 1,910 goal-conflict publications were initially identified; following review, 8 simple avoidance and 11 goal-conflict studies were included, with 5 datasets used in a preliminary meta-analysis. A move from forebrain-to-midbrain activation as threat becomes more pertinent was noted, indicating support for the Reinforcement Sensitivity Theory of behaviour and general compatibility with animal work. However, these findings were not reflected in the subsequent preliminary meta-analysis. This review highlights the considerable heterogeneity in currently available defensive behaviour paradigms and the lack of research in clinically relevant populations.

Keywords

Neuroimaging, human behaviour, anxiety, threat response, goal-conflict, individual differences

Background

Defensive behaviour and abnormal sensitivity to threat has been linked to psychopathology (Bijttebier, Beck, Claes, & Vandereycken, 2009; Mitchell et al., 2007), with particular relevance to anxiety disorders (McNaughton & Corr, 2004). Avoidance of threat (active movement away from threat) and approach-avoidance during goal-conflict (movements or decision making designed to collect information about a situation, or move towards reward when there is a risk of adverse event) are key aspects of human defensive behaviour (Hundt, Nelson-Gray, Kimbrel, Mitchell, & Kwapil, 2007; McNaughton & Corr,

2004). These behaviours form systems that are considered orthogonal but functionally interdependent, forming behavioural response to threat and threat-reward conflict (Jackson, 2009; McNaughton & Corr, 2004). In addition to an association with mental health pathology defensive behaviours are thought to show considerable individual differences (Corr, 2013).

Despite the clinical relevance of defensive behaviours, to date much of the experimental work used animal models (Kirlic, Young, & Aupperle, 2017). Rodent research involves a range of established avoidance/approach-avoidance tasks, from exploratory behaviour tasks such as the elevated plus maze (Pellow, Chopin, File, & Briley, 1985) to those using punishment for induction of conflict such as the Vogel conflict test (Vogel, Beer, & Clody, 1971), depicting neural activation in non-human animals (Davis, Walker, Miles, & Grillon, 2010; Grillon, Morgan, Davis, & Southwick, 1998; Kumar, Bhat, & Kumar, 2013). The rodent work has highlighted the amygdala and hippocampus (Choi & Kim, 2010; Kirlic et al., 2017; Möller, Wiklund, Sommer, Thorsell, & Heilig, 1997), periaqueductal grey (PAG) and midbrain (Fanselow, 1994) in response to threat and threatening conflict. Though animal findings are comprehensive and largely consistent, replication in humans has sometimes been problematic (Blanchard, 2017; Corr, 2002). As such, a review of the available literature concerning defensive behaviour in humans is both timely and may provide direction for future research.

Reinforcement Sensitivity Theory (RST) outlines human defensive behaviour, separating simple avoidance and goal-conflict both behaviourally and clinically and predicting involvement of specific neural regions, with activation progressing from cortical to subcortical as threat increases (see Figure 1)(McNaughton & Corr, 2004). Maladaptive avoidance of threat is characteristic of panic and phobic disorders (Blanchard, Hynd, Minke, Minemoto, & Blanchard, 2001; McNaughton & Corr, 2004) and abnormal response to conflicting stimuli is linked to diagnoses such as generalized anxiety disorder (Bijttebier et

al.,

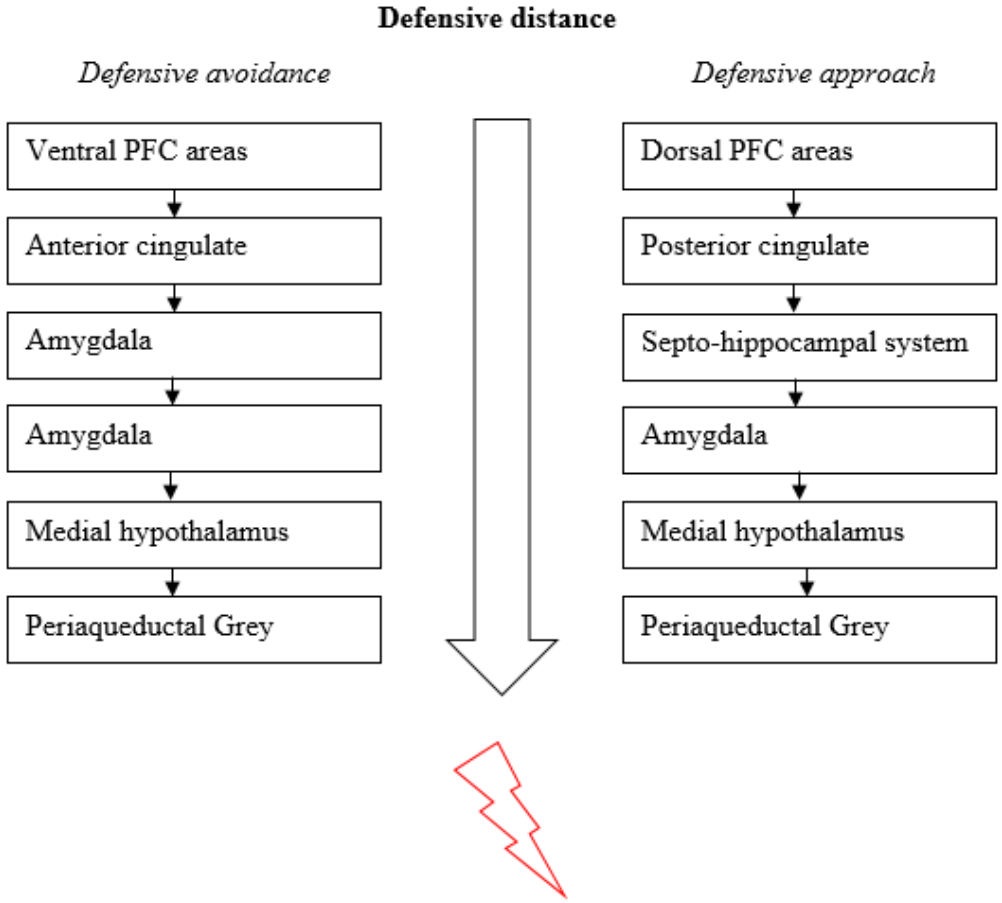

Figure 1 Neural activation during defensive behaviour under increasing threat levels, based on McNaughton & Corr (2004)

2009; Hundt et al., 2007; Kasch, Rottenberg, Arnow, & Gotlib, 2002); therefore understanding of human defensive behaviour architecture is vital to diagnosis and treatment (LeDoux, Moscarello, Sears, & Campese, 2017). As neuroimaging becomes increasingly prevalent and more sophisticated, development of accurate tools to understand neural correlates of behaviour is crucial. Despite simple avoidance and goal-conflict behaviour having clear clinical relevance (McNaughton & Corr, 2004; McNaughton & Gray, 2000) and being well documented in animal models (Blanchard et al., 2001; Kirlic et al., 2017; Kumar et al., 2013) with a clear human neural hypothesis (Kirlic et al., 2017; McNaughton & Corr, 2004), a systematic review or meta-analysis of the evidence regarding neural systems involved in human defensive behaviour has not yet been conducted.

Aims & hypothesis

This review explores the available human functional imaging paradigms for exploration of threat-related behaviours, and synthesises the neural activation reported by individual studies. Given the clinical relevance of defensive behaviour and the paucity of a synthesized body of human translational work, the aim of this review is three-fold: (1) to provide an overview of neural activation via human functional magnetic resonance imaging (fMRI), Positron Emission Tomography (PET), Magnetoencephalography (MEG) or Single Positron Emission Tomography (SPECT) of active behavioural goal-conflict and simple avoidance tasks, (2) outline physiological/self-report measures as validation of findings; and (3) assuming sufficient homogeneity and data, a meta-analysis of neural activation to provide additional insight. It is expected that the neural predictions of RST and animal work will be supported in this review.

Materials & methods

Literature search

Literature searches for English language papers were conducted using Scopus (Elsevier; www.scopus.com), PsychInfo (American Psychological Association; accessed via Ovid Technologies Inc, www.ovid.com), and Web of Science (Clarivate Analytics; www.webofknowledge.com). Search results were extracted March 2018, with no date limiters. Titles and abstracts were assessed, with those appropriate undergoing full text review. Reference lists were manually checked for additional studies. The search terms were chosen to identify goal-conflict approach-avoid tasks and threat avoidance behavioural tools in studies using imaging techniques, excluding lesion studies. Search terms were as follows: (“threat avoid*" OR ("threat" AND "avoid*") OR "defensive r*" OR "fight flight and freeze system" OR "fight" OR "flight" OR "freeze" OR "FFFS" OR "behavio* avoid*") AND (

"*MRI" OR "SPECT" OR "PET" OR "*magnetic resonance imaging") AND ("threat" OR "predator" OR "fear" OR "anxiety").

Outcome measures

Region of Interest (ROI) data or whole-brain derived data was accepted for systematic review. To prevent bias, whole-brain data or Statistical Parametric Mapping (SPM) t-maps were required for the meta-analysis. Articles that used a regions of interest (ROI) only, did not apply consistent statistical thresholds throughout the brain, or did not report peak coordinates in stereotactic space were excluded from the meta-analysis. The authors of work selected for meta-analysis were contacted requesting whole-brain or t-map data.

Study selection

Titles, authorships and abstracts were downloaded and formatted in to an excel document. Duplicates were manually removed. One author screened the titles and abstracts of all non-duplicate items, excluding the ineligible articles. Two authors assessed the eligibility of potential inclusions, reaching 100% agreement.

Inclusion & exclusion criteria

Studies were included if they met the following criteria. (1) Primary studies, exploring approach-avoidance (goal-conflict) or simple avoidance active behavioural task in presence of threat/risk of threat (including physical punishment and loss of accrued prizes) through active response (including pre-programmed outcomes, providing individuals are unaware) as passive viewing of stimuli do not have direct implications for avoidance/approach (Kirlic et al., 2017); (2) Clear description of the activation interaction presented. (3) Written in English. (4) Involved adult samples (≥18 years of age). (5) Samples were either healthy controls and/or anxiety diagnosed.

Studies were excluded if they involved samples with neurodevelopmental, neurodegenerative or lesion-based conditions, though healthy control arm were included. Psychiatric conditions beyond those outlined above are basis for exclusion, except for concomitant depression due to the high co-morbidity between these conditions. Studies recruiting individuals with single- or main- diagnosis of depression were not included. Depression was considered outside the scope of this review; there is evidence to suggest that the association between threat sensitivity and anxiety is stronger than in depression (Naragon-Gainey, 2010).It may be that passive avoidance is more relevant to depression (Ferster, 1973; Ottenbreit & Dobson, 2004); further, though there is little research in to defensive behaviours in depression using active avoidance paradigms, one study using this approach found no significant differences in neural activation between depressed patients and healthy controls (Marwood, 2017). Methodology outside of fMRI, MRI, PET, MEG and SPECT will be excluded as beyond the scope of this review.

Seed-based d Mapping (SDM) meta-analysis

SDM is a well validated (Radua et al., 2012) meta-analytic technique using a voxel-based approach. SDM uses whole-brain co-ordinates or SPM t-maps to calculate effect sizes from each included study, weighted by sample size to account for variance between studies. It has strict criteria for data inclusion such as excluding studies which do not report whole brain results to reduce publication bias. The SDM software package is available for free online (www.sdmproject.com). Analysis was conducted with SDM v5.15. Our analysis was thresholded at $p < .005$ and discarded clusters with voxels <10 to reduce risk of false positives, in line with other SDM based reviews (Radua & Mataix-Cols, 2012). Five studies were included (2 using whole-brain co-ordinates, 3 using an SPM t-map). Contrasts were of simple avoidance of threat only, as suitable heterogeneity and power was not possible within

the goal-conflict grouping; details of all included are available in table 5. Data sets that did not provide t-statistics were converted using the SDM package.

Results

Literature search

The search criteria identified 1,910 goal-conflict (Scopus, n = 1733; PsychInfo, n = 51; Web of Science, n = 25, unique) and 1,348 simple avoidance (Scopus, n = 1083; PsychInfo, n = 115; Web of Science, n = 150, unique) articles. One further article was identified through reference lists. Full text review was conducted on 12 simple-avoidance and 11 goal-conflict studies; four simple-avoidance studies were excluded as this stage due to the paradigm involving passive avoidance only (i.e. participants could not actively respond of their own will in order to promote/prevent avoidance), in line with our exclusion criteria. After full text review, 11 goal-conflict and eight simple-avoidance experimental papers were included. See

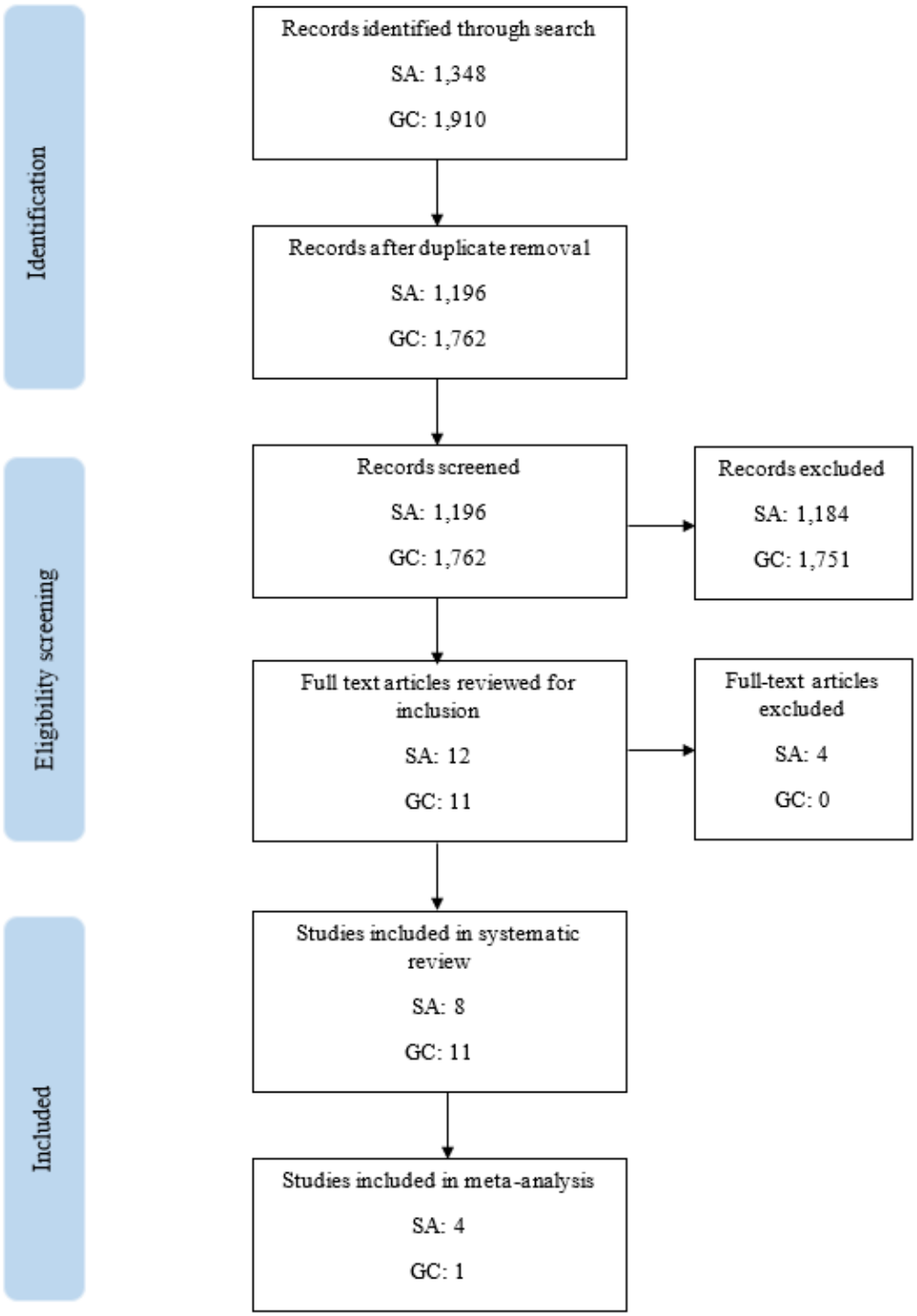

Figure 2 for flowchart of selection process.

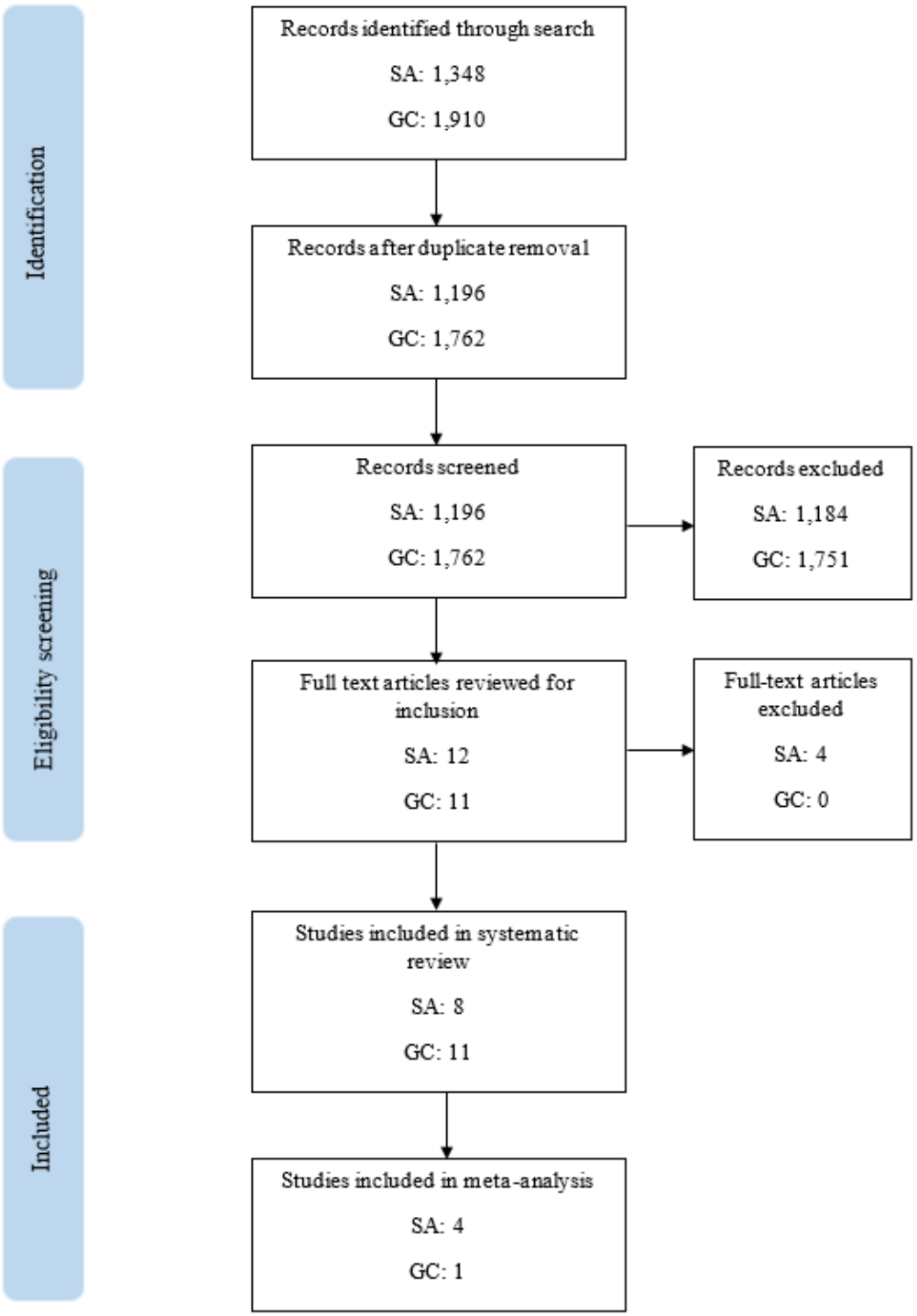

Figure 2 PRISMA diagram of included studies. Abbreviations: SA, Simple Avoidance; GC, Goal-Conflict

Description of selected studies

The identified paradigms were highly diverse. Simple avoidance tasks included: (1) maze/pathway tasks with virtual predators, n = 3; (2) non-chase response to prevent aversive event, n = 5. See table 1 for overview of studies. Goal-conflict were categorized as: (1) maze/open space/runway tasks with virtual predator, n = 5; (2) response (option selection) to prevent/encourage event tasks, n = 6. See table 2 for overview of studies. All studies used healthy controls only, and three included pharmacology.

Threat stimuli. Simple avoidance used two types of threat stimuli, physical threat (electric shock, n = 5, loud noise, n = 2) and loss of tokens/prizes (n = 1). Goal-conflict trials tended towards token/prize loss (n = 6), but also used physical threat (n = 2) and aversive images (n = 4).

Goal-conflict rewarding stimuli. Token/prize gain was the most frequent rewarding stimulus (n = 8), though pleasant images were also used (n = 3).

Simple avoidance tasks

Table 1 details the design of included studies. Defensive distance (i.e. distance from threat), threat anticipation (activation during threat cueing), reception of aversive outcome, and the level of threat presented were varyingly controlled. The latter was manipulated through predetermined probability of capture (e.g. Montoya, van Honk, Bos, & Terburg, 2015; Schlund et al., 2016; Wendt, Löw, Weymar, Lotze, & Hamm, 2017) stratified predator strength (Mobbs, Petrovic, Marchant, Hassabis, & Weiskopf, 2007; Mobbs et al., 2009) or in one case, visibility of predator (Rigoli, Ewbank, Dalgleish, & Calder, 2016). Trials using spatial navigation and an unpredictability of predator-threat are similar to rodent models, such as the Mouse Defence Test Battery (Blanchard, Griebel, & Blanchard, 2003). Fear

conditioning was used in a number of studies, requiring implicit learning of behaviour allowing threat avoidance (Boeke, Moscarello, LeDoux, Phelps, & Hartley, 2017; Collins, Mendelsohn, Cain, & Schiller, 2014; Schlund et al., 2016). Interestingly, only one trial permitted 'freezing' behaviour (Wendt et al., 2017). One trial had a pharmacological approach, exploring the role of cortisol in defensive behaviour (Montoya et al., 2015).

Simple avoidance: neural activation.

Table 3 shows neural activations in simple avoidance tasks. The key finding was a change from forebrain-to-midbrain activation as the threat came closer; specifically, activation changes from prefrontal cortices to the periaqueductal grey (PAG) and midbrain (Mobbs et al., 2007, 2009; Montoya et al., 2015; Wendt et al., 2017).

Prefrontal areas. Increased activity in prefrontal areas (ventromedial, dorsolateral and dorsomedial PFC) and cingulate cortices (CC; anterior and/or posterior) was observed in response to threat presence generally (Collins et al., 2014; Mobbs et al., 2009; Montoya et al., 2015; Schlund et al., 2016). Specifically, activation in these areas was associated with distal (Mobbs et al., 2009; Mobbs et al., 2007; Wendt et al., 2017), unavoidable (relative to avoidable) (Montoya et al., 2015; Schlund et al., 2016), or hidden (relative to visible) (Rigoli, Pavone, & Pezzulo, 2012) threat. Heightened activity in these areas were commonly associated with high threat levels (Boeke et al., 2017; Collins et al., 2014; Mobbs et al., 2007, 2009; Montoya et al., 2015; Wendt et al., 2017), though occasionally dorsal medial PFC area activation was present in low/absent threat situations (Collins et al., 2014; Mobbs et al., 2007). However, a handful of these paradigms also highlighted the anterior CC in proximal threat (Mobbs et al., 2007, 2009). Ventromedial PFC activation was shown to correlate with decreased locomotor errors during escape from threat, in one study (Mobbs et al., 2009).

Brain activation and threat

PAG & midbrain. PAG and midbrain areas were shown to activate in response to threat presence (Boeke et al., 2017; Collins et al., 2014). In contrast to the PFC, heightened activation in the PAG and midbrain areas was observed when threat was proximal and/or high (Mobbs et al., 2007, 2009; Wendt et al., 2017), or visible (relative to invisible; Rigoli et al., 2012). Complementing prefrontal data, PAG and midbrain activation were linked to increased locomotive errors (Mobbs et al., 2009); though somewhat contradictorily the anterior CC was also engaged. One trial suggested that midbrain activation in response to threat may be modulated by cortisol levels (Montoya et al., 2015).

Insula cortex. Anterior insula activation was associated with presence of threat (Collins et al., 2014; Montoya et al., 2015), with some evidence of differential posterior activation in threat-absent trials (Collins et al., 2014). This dual purpose was also present in regards to defensive distance, with both proximal (Mobbs et al., 2009; Wendt et al., 2017) and distal (Wendt et al., 2017) threats leading to heightened BOLD response. Anterior insula activation was linked to increased errors during threat exposure (Mobbs et al., 2009), anticipation of threat (Montoya et al., 2015) and reception of aversive stimuli due to non-avoidance of threat (Wendt et al., 2017).

Limbic system. Unsurprisingly, the amygdala was shown to respond to presence of threat, whether proximal (Mobbs et al., 2007, 2009; Wendt et al., 2017), distal (Mobbs et al., 2009) or hidden (Rigoli et al., 2012). Dorsal amygdala function was specifically linked to threat proximity, whilst basolateral amygdala (BLA) function was associated with distal threat (Mobbs et al., 2007), though other studies have found the direction of this finding to be variable (Montoya et al., 2015; Wendt et al., 2017), potentially due to issues in subdividing the amygdala. Thalamus (posterior and mediodorsal) and hypothalamus activation were linked to threat presence, regardless of distance from the subject, though the hippocampal structures activated only in response to distal (Mobbs et al., 2009) and high (Mobbs et al.,

2007) threat levels. The hippocampus also reacted during exposure to hidden threat (Rigoli et al., 2012), and curiously during non-threat exposure in one trial (Collins et al., 2014).

Goal-conflict tasks

The types of goal-conflict task were similar to simple avoidance categorisation. A subset of the navigation paradigms limited response to a restricted runway, with participants able to show level of approach/avoidance behaviour by placement along it (Aupperle, Melrose, Francisco, Paulus, & Stein, 2015; Schlund et al., 2016). In tasks requiring option selection to indicate response, 3 paradigms used pressure-sensitive joysticks to indicate choice (Cunningham, Arbuckle, Jahn, Mowrer, & Abduljalil, 2011; Radke et al., 2017, 2015). Threat (and/or reward) level was manipulated as in simple avoidance. This was via stratified threat/reward pairings (Loh et al., 2016; Schlund et al., 2016), probability of threat/reward (Bach et al., 2014; Khemka, Barnes, Dolan, & Bach, 2017; Talmi, Dayan, Kiebel, Frith, & Dolan, 2009) or both (Aupperle et al., 2015; Gonen et al., 2016). A number of studies removed choice, telling participants which action to use (Cunningham et al., 2011; Radke et al., 2017, 2015). One paradigm removed in-trial feedback entirely, presenting only stimuli conditioned as threat or reward representations (O'Neil et al., 2015). Pharmacological intervention featured in two paradigms by the same authors, one exploring testosterone (Radke et al., 2015) and the other oxytocin (Radke et al., 2017).

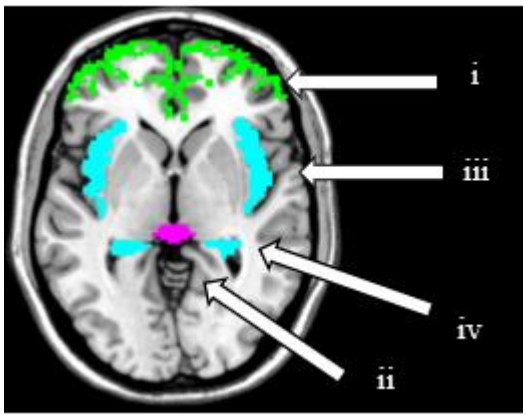

Figure 3 Regions reported in systematic review of human defensive behaviour: i) ventromedial and dorsolateral prefrontal cortices showing higher activation in response to distal threats; ii) midbrain and periaqueductal grey regions showing higher activation in response to proximal threats; iii) insular cortices, activating in response to conflict and threat; iv) hippocampus and amygdala (posterior hippocampus shown), the amygdala activates in response to threat, and hippocampus, showing activation in simple avoidance and goal-conflict trial types, though when motion is controlled for hippocampal activation appears relevant to goal-conflict trials only

Goal-conflict, neural activation

Neural activations are displayed in table 4. Paradigms compared conflict and non-conflict situations, varied threat/reward level and assessed activation associated with motivational direction.

Prefrontal areas. Prefrontal and CC activation was associated with conflict (Aupperle et al., 2015; Radke et al., 2017, 2015), though not necessarily high conflict (Gonen et al., 2016). PFC activation was also associated with errors during reward prediction (Talmi et al., 2009) and activated during decision-making generally (Loh et al., 2016), suggesting an assessment role for these areas. A point of separation is observed in PFC activation during

large value-outcome (reward vs threat) differences, and CC activation in response to small value-outcome differences (Schlund et al., 2016).

PAG and midbrain regions. Relative to threat avoidance, the PAG and midbrain regions featured less here. PAG activation was activated in high conflict (Gonen et al., 2016), as were the putamen, caudate and thalamus (Aupperle et al., 2015; O'Neil et al., 2015).

Insula cortex. Insula activity was widely relevant, indicated in response to conflict (Aupperle et al., 2015), increasing threat (Bach et al., 2014), errors in reward prediction (Talmi et al., 2009), choice selection (Loh et al., 2016), approach (relative to avoidance) (O'Neil et al., 2015) and small value-outcome differences (Schlund et al., 2016). Conversely, this area also activated in response to non-conflict scenarios (Aupperle et al., 2015).

Limbic system. Activation of the hippocampus (and parahippocampal gyri) was associated with conflict (O'Neil et al., 2015), decision making (Loh et al., 2016), increased threat (Bach et al., 2014), approach ((O'Neil et al., 2015) and successful avoidance of loss (Loh et al., 2016). One study used a ROI MEG approach, identifying increased hippocampal oscillation in the right, and decreased in the left, hemisphere during high threat conflict (Khemka et al., 2017). Hippocampal activation was also associated with threshold of outcome-values, when the value-difference between outcomes is smallest (Schlund et al., 2016). As in simple avoidance, amygdala activation was raised in high threat scenarios (Bach et al., 2014), as well as during conflict (O'Neil et al., 2015). Stimuli and motivational valence was also linked to the amygdala, with increased activity in response to emotional stimuli (Cunningham et al., 2011) and approach behaviour (Cunningham et al., 2011; Radke et al., 2017), the latter reflective of activation during proximal threat outlined in simple avoidance (Mobbs et al., 2007, 2009). One trial indicated testosterone in amygdala response, increasing activation during approach and decreasing during avoidance (Radke et al., 2015). Similarly,

oxytocin may play a role, as indicated by amygdala deactivation during approach, though not avoidance (Radke et al., 2017).

Self-report and physiological data

Not all paradigms used self-report or physiological data as validation of threat/reward experience. Skin conductance response (SCR) (Gonen et al., 2016; Mobbs et al., 2009; Schlund et al., 2016; Talmi et al., 2009; Wendt et al., 2017) or self-report anxious traits (Collins et al., 2014; Loh et al., 2017; Mobbs et al., 2009) were used to support threat value, though not always successfully (Aupperle et al., 2015; Boeke et al., 2017; Radke et al., 2017; Rigoli et al., 2016). Brain activity was also validated with self-report, as shown in the association between PAG activation and self-reported dread (Mobbs et al., 2007) and amygdala response and trait anxiety (Mobbs et al., 2009), during threat exposure. Individual differences in personality were also shown to have an impact on threat sensitivity, as represented by differences in VTA and VS activation during approach (Gonen et al., 2016), and interactions between neuroticism and amygdala response to threat (Cunningham et al., 2011). See tables 3 and 4 for further details.

Preliminary meta-analysis of simple avoidance

As heterogeneity between tasks was high, five studies with comparable designs were included in a meta-analysis of avoidance of threat (n = 151 healthy control participants across the 5 independent publications). These studies were included based on similarity of contrasts analysed (i.e. all studies included analysis comparing avoidance in high vs. low or absent-threat conditions). Of the 8 simple-avoidance studies included in the systematic review, 4 were not included in the meta-analysis (due to fundamental difference in contrasts, i.e. not directly comparing high vs. low/absent threat, n = 2; and due to availability of data, i.e. no whole-brain co-ordinates or t-maps available, n = 2). One goal-conflict study (Bach et al., 2014) included a simple-avoidance contrast analysed separately (from the goal-conflict

analysis), and so this contrast was also included in the simple-avoidance meta-analysis. A meta-analysis of neural activation in goal-conflict was not possible, due to study design and contrast analysis heterogeneity. Studies using Montreal Neurological Institute (MNI), Talarach and FMRIB Software Library (FSL) standardised space and providing either whole-brain co-ordinates or SPM statistical maps were included. See table 5 for details of studies and contrasts.

As shown in table 6, several common brain regions were identified, mostly centred on the frontal gyri. Jack-knife sensitivity analysis was conducted to assess robustness; no finding was reliably present across all five studies (table 6). All areas driven by ≥4 studies are indicated in the table as particularly robust. Activation of the right medial prefrontal cortex region (middle frontal gyrus) but deactivation of the left middle frontal gyrus was present in all but 1 study, as indicated in table 6. In contrast to the systematic review, midbrain activation during greater threat and forebrain activation in lower threat was not shown in this preliminary meta-analysis. Funnel plots were created for all regions driven by ≥4 studies. The funnel plots appeared well distributed and all Eggar's bias tests were non-significant suggesting minimal bias. However, caution is advised when interpreting plots with small numbers of studies. Too few studies were available to perform meta-regression to assess heterogeneity (Radua, van den Heuvel, & Surguladze, 2010) between response types (button press vs. avatar movement). **Error! Reference source not found.** (MRIcron; www.nitrc.org) shows activations and deactivations relating to threat level prior to jack-knife analysis.

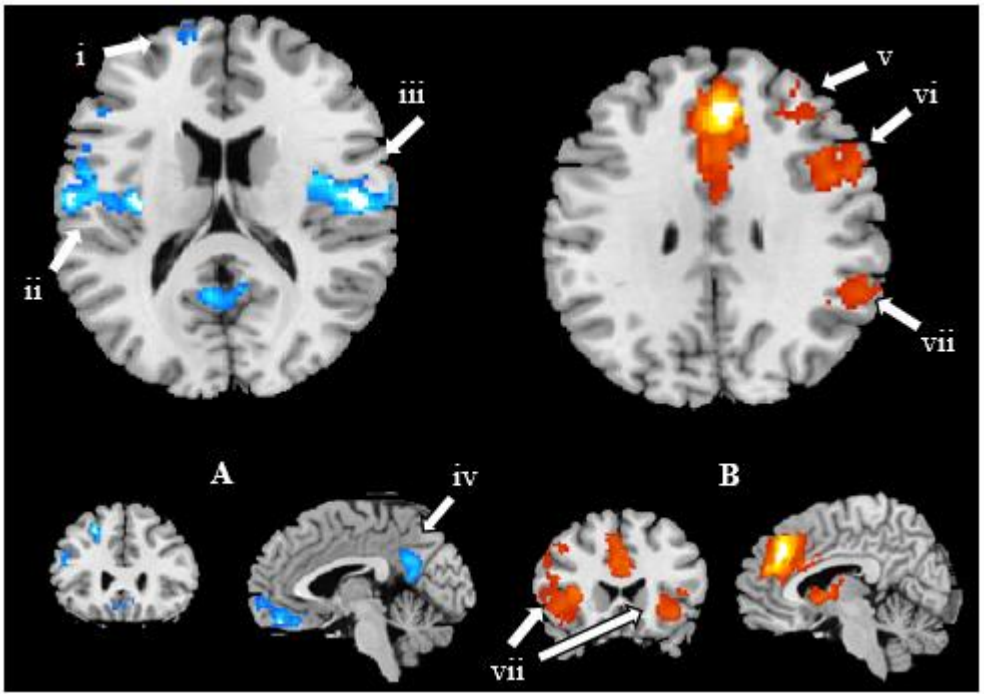

Figure 4 Brain regions identified in avoidance in threat vs. low or non-threat. A) Deactivations in high vs. low threat (SDM threshold regions of > 10 voxels, peak Z value = 3.1); B) Activations in high vs. low threat (SDM threshold regions of >10 voxels, peak negative Z value = -2.130. Deactivations (in threat vs. low/non-threat) are indicated in the left-side panel (A), with shades of blue reflective of intensity of deactivation in the: i) left superior frontal gyrus (-14, 62, 16); ii) left superior temporal (-60, -12, 10) and middle frontal gyri (-22, 26, 44); iii) right temporal pole (40, 6, -22)/fronto-insula tract (52, -10, 18); iv) left precuneus (-10, -54, 30). Activations (in threat vs. low/non-threat) are shown in the right-side panel (B): v) right superior frontal gyrus, dorsolateral (22, 42, 38) & medial (4, 38, 38); vi) right middle gyrus (40, 28, 36); vii) right supramarginal gyrus (56, -44, 32); viii) left (-38, 16, 10) & right (44, -48, 36) superior longitudinal fasciculus.

Discussion

This review explored neural activation in human defensive reactions. Simple avoidance was characterised by a forebrain-to-midbrain change in activation as threat approaches (Mobbs et al., 2007, 2009; Montoya et al., 2015; Wendt et al., 2017), supported in the review via differential midbrain and forebrain activation (Montoya et al., 2015; Rigoli et al., 2012). A cortical-subcortical change is in line with animal work (Blanchard, 2017) and prominent theories of human defence (Corr, 2013; McNaughton & Corr, 2004). It also reflects the clinical literature, as shown in PFC activation in anxiety (Myers-Schulz &

Koenigs, 2012) and PAG association with panic (Canteras & Graeff, 2014) (see Figure 1). Observed PFC and PAG functional connectivity (Chan et al., 2011) supports the concept of interactive suppression dependent on threat proximity (Mobbs et al., 2007).

More innate bottom-up processing may lead in situations requiring fast response to imminent threat, with higher-order processing associated with evaluative responses to proximal threat and decision making. In support, conflict and decision-making was linked to high-order areas such as the PFC and CC in goal-conflict trials (Aupperle et al, 2015; Loh et al., 2017) and more accurate escape behaviour (Mobbs et al., 2009), as well as difficult-to-gauge threat, whilst PAG activation was shown in response to clear threat (Rigoli et al., 2012). Change in activation from the dorsal PFC through posterior cingulate, septo-hippocampal system, the amygdala, the medial hypothalamus to the PAG during increasing defensive approach is also predicted in the literature (McNaughton & Corr, 2004). This review provides some support for this, as threat-approach and exploration was also shown to involve the PAG, midbrain and limbic areas (Bach et al., 2014; Cunningham et al., 2011; O'Neil et al., 2015). The pattern of activation revealed by the systematic review is reminiscent of the default mode network (DMN), with both the CC and PFC key components (Greicius, Krasnow, Reiss, & Menon, 2003). There is evidence to suggest the DMN may be altered in individuals with anxiety disorder (Zhao et al., 2007), supporting the concept of these regions as integral to anxiety-related neural circuitry and behaviour. Further, the DMN has been proposed as integral to neuroticism and self-generated though processes (Adam M. Perkins, Arnone, Smallwood, & Mobbs, 2015), including cognitive processes such as worry and rumination, which are integral to affective disorders (American Psychiatric Association, 2013).

Conflict was linked to both subcortical (insula, midbrain, PAG, hippocampus and amygdala; Aupperle et al., 2015; Bach et al., 2014; Gonen et al., 2016; Khemka et al., 2017)

and cortical (PFC and CC; Gonen et al., 2016) activation. Frontal activation is as expected, considering the role of PFC and CC in conflict monitoring (Botvinick, 2007) and executive functioning (Koechlin & Summerfield, 2007). Interestingly, PFC response to conflict is modulated by testosterone (Radke et al., 2015), with clinical implications for maladaptive approach behaviours; as the tested sample were all male, no comment can be made on how gender differences may feature in this relationship. Theories hold a tentative role for the (anterior) CC in conflict resolution in defensive behaviour, and (particularly the dorsal aspect of) both the CC and PFC (McNaughton & Corr, 2004), which is supported here. One study stratified conflict, indicating higher conflict was mostly associated with subcortical regions, and lower conflict with cortical (Gonen et al., 2016), reflective of threat proximity findings in simple avoidance. However there is far from a consensus as a number of studies report PFC and/or CC activation during absence of conflict (Aupperle et al., 2015; Gonen et al., 2016) suggesting further work is required.

The insula activation was shown in both simple avoidance and goal-conflict. The insula has previously been associated with conflict (Roberts & Hall, 2008) stimuli salience (Stein & Paulus, 2009), processing of pain and bodily sensation (Kirlic et al., 2017; Talmi et al., 2009), potentially relaying to the amygdala (Phelps et al., 2001). Seeley et al identified a salience-value processing network including the anterior insula, amygdala and dorsal/anterior CC (Seeley et al., 2007). Anterior CC, anterior insula and inferior frontal regions activation was observed when difference between threat and reward outcomes was increased (i.e. both outcomes have high salience), and ventromedial and dorsolateral PFC activation when decreased (Schlund et al., 2016). Frontal-cortical regions and the CC have been shown to activate in low goal-conflict (Gonen et al., 2016), and conflict present vs. absent situations (Aupperle et al., 2015; O'Neil et al., 2015) though this activation was not exaggerated in higher goal-conflict. The VTA and VS showed increased activity in higher conflict (Gonen et

al., 2016), in line with animal work suggesting a role for these areas in motivational response (Haber & Knutson, 2010; Williams, Rolls, Leonard, & Stern, 1993). The reward system is pertinent here, considering the role of the OFC, anterior CC, VS and amygdala identified in this review, and within the reward-circuitry of the human brain (Haber & Knutson, 2010). Deactivation of the right fronto-insular tract was noted in the meta-analysis; the insula is typically involved in processing of conflict and bodily sensation/pain (Kirlic et al., 2017; Roberts & Hall, 2008), deactivation of its connections with frontal regions could reflect deactivation of frontal processing in high threat due to reliance on innate bottom-up processing.

The hippocampus is considered integral to approach-avoidance conflict specifically (Ito & Lee, 2016; Perkins et al., 2013), as it is linked to sustained anxiety rather than fear (McNaughton & Corr, 2004). Anxiety is likely related to more distal and unpredictable fears (Davis et al., 2010), reflective of the hippocampal activity observed here. A key role of the hippocampus is spatial function and memory (Eichenbaum & Cohen, 2014). Hippocampal involvement in simple avoidance was present in tasks involving a high degree of spatial functioning, whether navigation within a 'maze'(Mobbs et al., 2009), prediction of spatial location of an invisible threat (Rigoli et al., 2016), or specific spatial orientation (Collins et al., 2014). Goal-conflict trials involving hippocampal activation did include some spatial processing (Bach et al., 2014; Khemka et al., 2017; Schlund et al., 2016), but unlike simple avoidance activation was also observed in goal-conflict without spatial demands (Loh et al., 2016; O'Neil et al., 2015; Talmi et al., 2009). These findings support a distinct role for the hippocampus in goal-conflict beyond spatial processing, in line with prominent theories (McNaughton & Corr, 2004) and animal work (Kirlic et al., 2017).

Amygdala activation was observed in response to threat (Mobbs et al., 2007, 2009), as in animal models (Blanchard, 2017; Davis et al., 2010). Amygdala response in defensive

distance is thought to divide across both anticipation and avoidance of threat (see figure 1) (Canteras & Graeff, 2014). Research has distinguished basolateral amygdala (BLA) and the central amygdala (CeA)/basal nucleus of the stria terminalis (BNST) activation, though these regions are highly interconnected (Davis et al., 2010; Dong, Petrovich, & Swanson, 2001). The BLA is associated with threat value judgement via connections with the ventromedial PFC and OFC and the CeA/BNST with behavioural and basic autonomic system activity via the PAG (Fanselow, 1994; Mobbs et al., 2009; Quirk, Likhtik, Pelletier, & Paré, 2003). Animal work indicates the CeA and the BNST are distinct, contributing to fear and anxiety respectively (Kumar et al., 2013), with improved avoidance behaviour after CeA lesions (LeDoux et al., 2017). Whilst human amygdala lesions are associated with reduced fear behaviour in response to threat (Korn et al., 2016), lesions restricted to the BLA have been linked to fear hypervigilance (Terburg et al., 2012). Though not distinguished in goal-conflict paradigms, one simple avoidance experiment associated the BLA and CeA with low and high threat, respectively (Mobbs et al., 2009). The BLA in particular has been proposed as a factor in the cortical-to-subcortical activation change in fear responding through its inhibitory role (Terburg et al., 2012). Activation of regions such as the CeA and BNST are of clinical interest, given evidence that N-methyl-D-aspartate receptor function facilitation can increase context-specific extinction (Perusini & Fanselow, 2015; Walker & Davis, 2002), with clear implications for treatment. Similarly, separation of anterior and posterior cingulate cortical regions has been hypothesised representing defensive avoidance and defensive approach at roughly equivalent distances (McNaughton & Corr, 2004), though this is not supported in this review.

Though the systematic review supports RST and animal work, the meta-analysis did not clearly support these findings. Prefrontal areas accounted for the most robust findings, such as activation of the middle frontal gyrus and superior frontal gyrus was shown in high

threat relative to low/absent threat situations. This was unexpected, given the link between prefrontal regions and lower threat levels supported in the systematic review. However, this is a restricted portion of the PFC and heterogeneity of threat type and task behaviour should be considered when interpreting this finding. Activation of the superior frontal gyri has been linked to attention and attention shifting (Nagahama et al., 1999), suggesting this region could be linked to general behaviour during active tasks of any description beyond the involvement of threat. Lateralization of region activation was apparent in the meta-analysis output, with high threat relating to greater activation of mostly right hemisphere regions and deactivation of the left side broadly speaking. This is in line with previous work linking the right hemisphere specifically to avoidant behaviour (Aupperle et al., 2015; Kirlic et al., 2017). Absence of activation of the PAG, midbrain, amygdala and hippocampus despite presence in individual studies and a focus on these areas in ROI analysis is notable. There are several important caveats to these findings: the number of studies included in the meta-analysis was small (Radua et al., 2012), making aspects of analysis potentially problematic (Radua et al., 2010), and the paradigms included maintained key differences (including consolidation of low-threat and non-threat against high threat). As such, further work with bigger samples and more consistent contrasts is required to build on this preliminary analysis.

Despite support for RST indicated in this review, self-report measures of RST systems (e.g. behavioural inhibition system/behavioural activation system response scale, BIS/BAS scale; Carver & White, 1994) were not shown to correlate with task behaviour or neural activation (Cunningham et al., 2011; Radke et al., 2015); however, the trials using these measures were largely restricted to ROI analysis, and the BIS/BAS scales were designed under older RST models, before separation of fear (simple avoidance) and anxiety (goal-conflict) as independent systems (Jackson, 2009; McNaughton & Corr, 2004). Compatibility of self-report and behavioural assessment of defensive reaction has previously

been questioned (LeDoux et al., 2017); comparison using updated self-report scales is a logical next step. A role for trait-anxiety was highlighted in simple avoidance (Collins et al., 2014; Mobbs et al., 2009; Rigoli et al., 2016), though only one goal-conflict trial found an association with behaviour (Loh et al., 2017). Both neuroticism and trait-anxiety were associated with amygdala and CC activity in response to threat (Cunningham et al., 2011; Mobbs et al., 2009). Individual sensitivity to pain was shown to attenuate reward seeking, associated with the OFC and cerebellum (Talmi et al., 2009), and individual difference in approach or avoidant personality traits was associated with differential VTA and VS (i.e. motivational response (Haber & Knutson, 2010) activation during conflict (Gonen et al., 2016). Though not entirely unanimous (e.g. (Aupperle et al., 2015), these findings highlight the importance of individual differences in understanding of both neural and behavioural defensive response, in line with contemporary views (Corr & Mobbs, 2018). As abnormal sensitivity to threat is considered a hallmark of anxiety disorders (Hundt et al., 2007; McNaughton & Corr, 2004) this is an important consideration. The accuracy of comparing self-report and behavioural measures is also raised, future work could compare neural activations presented here with neural activation during self-report measures of defensive behaviour.

Limitations

Due to the small number of studies with appropriate data available, the presented meta-analysis relies on a limited amount of data and as such must be interpreted with caution. Similarly, this meant a sensitivity analysis regarding stimuli type or comparing naturalistic vs. conditioned threat was not possible. Many trials consistently used pre-defined ROIs, presenting findings from only one- or two- regions. This approach potentially ignores the wealth of data available from elsewhere and may preclude unexpected activations. ROI selection is also at risk of a strong publication bias. Fear conditioning trials have a strong

presence in the translational literature (Kirlic et al., 2017) and were included within this review; future work would benefit from comparison of defensive behaviour in response to conditioned vs. naturalistic threat. The paradigms identified in this review show considerable variability in design. Despite identification of 19 studies, the heterogeneity of tasks in this field is so high that meta-analysis of data was restricted. However, heterogeneity can help identification of robust findings.

Future directions in study design

Several proposals are made for future paradigms. A paucity of work in anxious samples and its association with clinical understanding means prioritisation of comparative work between healthy and anxious samples is a priority; moving research beyond general anxiety is also of interest, as maladaptive defensive behaviours have also been highlighted as integral to disorders such as obsessive-compulsive disorder and autism (Gillan et al., 2014; Servatius, 2016). As aforementioned, neural activation during defensive behaviours in depression is unclear (Ferster, 1973; Marwood, 2017; Naragon-Gainey, 2010; Ottenbreit & Dobson, 2004), warranting further research within this diagnostic area. Only one study included freezing as a behavioural response, despite freezing being a core aspect of threat response system (FFFS) and common in rodent models. Development of paradigms able to support freezing as a legitimate response would enrich understanding of human fear. Given the association between experiments with high spatial components and hippocampal activation, care should be taken to avoid conflation; the use of joysticks to enlarge/shrink stimuli as a representation of approach/avoidance respectively (e.g. Radke et al., 2017, 2015;Cunningham et al., 2011) might be an alternative to maze/runway paradigms, though an assessment of potential spatial hippocampal involvement in this action is necessary first. However, maze paradigms remain faithful to the animal models that provided the basis for

the field and provide opportunity for simple avoidance and goal-conflict trials within one task which provides greater insight.

Given the unexpected activation of frontal regions during higher threat shown in the meta-analysis, and the link between frontal gyri and attention (Nagahama et al., 1999) the potential confound of attention level should be considered in future work. Though activation is believed to change from forebrain-to-midbrain with threat proximity, the nature of this change is unknown. Prolonged threat exposure with gradual proximity change would indicate whether the change in activation is a binary switch or a gradual change; identifying the turning point would be useful, especially if combined with a measure of individual difference such as neuroticism. Identification of different cut-off points associated with neuroticism score would be informative considering the link between neuroticism and risk of anxiety disorders (Lahey, 2009). The use of self-report and physiological data is inconsistent but recommended in future projects to represent participant experience of experimentally induced fear. In addition to neuroticism, state measures of anxiety or ongoing cardiac or skin conductance measures would be useful in tool validation (see Mobbs et al, 2007, 2009). There is evidence that alternate neural systems may be involved in the processing of monetary gain, relative to pain and affective threat outcomes (Kirlic et al., 2017). Some studies using monetary gain/reward also used physical threat stimuli such as electric shock (for example, Aupperle, Melrose, Francisco, Paulus, & Stein, 2015a; Talmi et al., 2009), causing potential confounds. The immediacy of outcome of these stimuli may also be an issue, with immediate shock/emotionally aversive imagery not necessarily equivalent to promise of money later. In future exploration of an immediate and physical reward stimuli such as pleasant smells or sweet drinks/food would be beneficial, as attempted by (Rzepa, Fisk, & McCabe, 2017).

Conclusion

Generally, a change from cortical to subcortical activation is observed in response to increasing threat, whether the threat is being avoided or approached, whilst conflict is associated with an array of cortical and subcortical regions. The findings are largely in line with the predictions of RST, basic reward circuitry and motivational salience. The findings are also supportive of the close extrapolation from animal to human work that has shaped the field. Hippocampal involvement in simple avoidance appears largely associated with spatial demands, distinct from its role in goal-conflict trials. A meta-analysis of threat avoidance neural activation did not indicate activation of the same regions as the systematic review, though the limitations of this analysis are highlighted. There is a dearth of exploration in anxious populations, despite a theoretical focus on links between clinical presentation and threat sensitivity. Understanding the neural circuitry underlying common anxiety-related behaviours is key to the development and refinement of treatments for psychiatric conditions caused by dysregulation in these regions. Several recommendations for future paradigms are outlined.

Acknowledgements and disclosures

Allan H. Young, Matthew Kempton, Steven C.R. Williams and Adam M. Perkins are supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. LM was supported by a joint Medical Research Council / Institute of Psychiatry, Psychology and Neuroscience PhD studentship. This report represents independent research funded by the NIHR Biomedical Research Centre and South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, NIHR or Department of Health.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Washington DC.

Aupperle, R. L., Melrose, A. J., Francisco, A., Paulus, M. P., & Stein, M. B. (2015). Neural substrates of approach-avoidance conflict decision-making. *Human Brain Mapping*, *36*(2), 449–462. https://doi.org/10.1002/hbm.22639

Bach, D. R., Guitart-Masip, M., Packard, P. A., Miró, J., Falip, M., Fuentemilla, L., & Dolan, R. J. (2014). Human hippocampus arbitrates approach-avoidance conflict. *Current Biology*, *24*(5), 541–547. https://doi.org/10.1016/j.cub.2014.01.046

Bijttebier, P., Beck, I., Claes, L., & Vandereycken, W. (2009). Gray's Reinforcement Sensitivity Theory as a framework for research on personality-psychopathology associations. *Clinical Psychology Review*, *29*(5), 421–430. https://doi.org/10.1016/j.cpr.2009.04.002

Blanchard, D. C. (2017). Translating dynamic defense patterns from rodents to people. *Neuroscience and Biobehavioral Reviews*, *76*, 22–28. https://doi.org/10.1016/j.neubiorev.2016.11.001

Blanchard, D. C., Griebel, G., & Blanchard, R. J. (2003). The Mouse Defense Test Battery: Pharmacological and behavioral assays for anxiety and panic. *European Journal of Pharmacology*, *463*(1–3), 97–116. https://doi.org/10.1016/S0014-2999(03)01276-7

Boeke, E. A., Moscarello, J. M., LeDoux, J. E., Phelps, E. A., & Hartley, C. A. (2017). Active Avoidance: Neural Mechanisms and Attenuation of Pavlovian Conditioned Responding. *The Journal of Neuroscience*, *37*(18), 4808–4818. https://doi.org/10.1523/JNEUROSCI.3261-16.2017

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience*, *7*(4), 356–366. https://doi.org/10.3758/CABN.7.4.356

Canteras, N. S., & Graeff, F. G. (2014). Executive and modulatory neural circuits of

defensive reactions: Implications for panic disorder. *Neuroscience and Biobehavioral Reviews*, *46*(P3), 352–364. https://doi.org/10.1016/j.neubiorev.2014.03.020

Caroline Blanchard, D., Hynd, A. L., Minke, K. A., Minemoto, T., & Blanchard, R. J. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. *Neuroscience & Biobehavioral Reviews*, *25*(7–8), 761–770. https://doi.org/10.1016/S0149-7634(01)00056-2

Carver, C. S., & White, T. L. (1994). Behavioral inhibition , behavioral activation , and affective responses to impending reward and punishment : The BIS / BAS Scales. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.67.2.319

Chan, T., Kyere, K., Davis, B. R., Shemyakin, A., Kabitzke, P. A., Shair, H. N., … Wiedenmayer, C. P. (2011). The Role of the Medial Prefrontal Cortex in Innate Fear Regulation in Infants, Juveniles, and Adolescents. *Journal of Neuroscience*, *31*(13), 4991–4999. https://doi.org/10.1523/JNEUROSCI.5216-10.2011

Choi, J. S., & Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proceedings of the National Academy of Sciences*, *107*(50), 21773–21777.

Collins, K. A., Mendelsohn, A., Cain, C. K., & Schiller, D. (2014). Taking Action in the Face of Threat: Neural Synchronization Predicts Adaptive Coping. *Journal of Neuroscience*, *34*(44), 14733–14738. https://doi.org/10.1523/JNEUROSCI.2152-14.2014

Corr, P. J. (2002). J.A. Gray's reinforcement sensitivity theory and frustrative nonreward: A theoretical note on expectancies in reactions to rewarding stimuli. *Personality and Individual Differences*, *32*(7), 1247–1253. https://doi.org/10.1016/S0191-8869(01)00115-5

Corr, P. J. (2013). Approach and avoidance behaviour: Multiple systems and their interactions. *Emotion Review*, *5*(3), 285–290. https://doi.org/10.1177/1754073913477507

Corr, P. J., & Mobbs, D. (2018). From epiphenomenon to biologically important phenomena. *Personality Neuroscience*, (May), 2016–2019. https://doi.org/https://doi.org/10.1017/pen.2017.1

Cunningham, W. A., Arbuckle, N. L., Jahn, A., Mowrer, S. M., & Abduljalil, A. M. (2011). Reprint of: Aspects of neuroticism and the amygdala: Chronic tuning from motivational styles. *Neuropsychologia*, *49*(4), 657–662. https://doi.org/10.1016/j.neuropsychologia.2011.02.027

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs Sustained Fear in Rats and Humans: Role of the Extended Amygdala in Fear vs Anxiety. *Neuropsychopharmacology*, *35*(1), 105–135. https://doi.org/10.1038/npp.2009.109

Dong, H. W., Petrovich, G. D., & Swanson, L. W. (2001). Topography of projections from amygdala to bed nuclei of the stria terminalis. *Brain Research Reviews*, *38*(1–2), 192–246. https://doi.org/10.1016/S0165-0173(01)00079-0

Eichenbaum, H., & Cohen, N. J. (2014). Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function? *Neuron*, *83*(4), 764–770. https://doi.org/10.1016/j.neuron.2014.07.032

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review*, *1*(4), 429–438. https://doi.org/10.3758/BF03210947

Ferster, C. B. (1973). A functional analysis of depression. *The American Psychologist*, *28*(10), 857–870.

Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., … Robbins, T. W. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological Psychiatry*, *75*(8), 631–638. https://doi.org/10.1016/j.biopsych.2013.02.002

Gonen, T., Soreq, E., Eldar, E., Ben-Simon, E., Raz, G., & Hendler, T. (2016). Human mesostriatal response tracks motivational tendencies under naturalistic goal conflict. *Social Cognitive and Affective Neuroscience*, *11*(6), 961–972. https://doi.org/10.1093/scan/nsw014

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(1), 253–258. https://doi.org/10.1073/pnas.0135058100

Grillon, C., Morgan, C. A., Davis, M., & Southwick, S. M. (1998). Effects of experimental

context and explicit threat cues on acoustic startle in Vietnam veterans with posttraumatic stress disorder. *Biological Psychiatry*, *44*(10), 1027–1036. https://doi.org/10.1016/S0006-3223(98)00034-1

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4–26. https://doi.org/10.1038/npp.2009.129

Hundt, N. E., Nelson-Gray, R. O., Kimbrel, N. A., Mitchell, J. T., & Kwapil, T. R. (2007). The interaction of reinforcement sensitivity and life events in the prediction of anhedonic depression and mixed anxiety-depression symptoms. *Personality and Individual Differences*, *43*(5), 1001–1012. https://doi.org/10.1016/j.paid.2007.02.021

Ito, R., & Lee, A. C. H. (2016). The role of the hippocampus in approach-avoidance conflict decision-making: Evidence from rodent and human studies. *Behavioural Brain Research*, *313*, 345–357. https://doi.org/10.1016/j.bbr.2016.07.039

Jackson, C. J. (2009). Jackson-5 scales of revised Reinforcement Sensitivity Theory (r-RST) and their application to dysfunctional real world outcomes. *Journal of Research in Personality*, *43*(4), 556–569. https://doi.org/10.1016/j.jrp.2009.02.007

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioural activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology*, *111*(4), 589.

Khemka, S., Barnes, G., Dolan, R. J., & Bach, D. R. (2017). Dissecting the Function of Hippocampal Oscillations in a Human Anxiety Model. *The Journal of Neuroscience*, *37*(29), 6869–6876. https://doi.org/10.1523/JNEUROSCI.1834-16.2017

Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translation paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy*, *96*, 14–29.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*(6), 229–235. https://doi.org/10.1016/j.tics.2007.04.005

Korn, C. W., Vunder, J., Mir??, J., Fuentemilla, L., Hurlemann, R., & Bach, D. R. (2016). Amygdala Lesions Reduce Anxiety-like Behavior in a Human Benzodiazepine-Sensitive Approach-Avoidance Conflict Test. *Biological Psychiatry*, *82*(7), 522–531.

https://doi.org/10.1016/j.biopsych.2017.01.018

Kumar, V., Bhat, Z. A., & Kumar, D. (2013). Animal models of anxiety: A comprehensive review. *Journal of Pharmacological and Toxicological Methods*, *68*(2), 175–183. https://doi.org/10.1016/J.VASCN.2013.05.003

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist*, *64*(4), 241–256. https://doi.org/10.1037/a0015309

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry*, *22*(1), 24–36. https://doi.org/10.1038/mp.2016.166

Loh, E., Kurth-Nelson, Z., Berron, D., Dayan, P., Duzel, E., Dolan, R., & Guitart-Masip, M. (2017). Parsing the Role of the Hippocampus in Approach-Avoidance Conflict. *Cerebral Cortex (New York, N.Y. : 1991)*, *27*(1), 201–215. https://doi.org/10.1093/cercor/bhw378

Marwood, L. (2017). *Novel functional imaging methods in depression and anxiety*. King's College London.

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews*, *28*(3), 285–305. https://doi.org/10.1016/j.neubiorev.2004.03.005

McNaughton, N., & Gray, J. A. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal of Affective Disorders*, *61*(3), 161–176. https://doi.org/10.1016/S0165-0327(00)00344-X

Mitchell, J. T., Kimbrel, N. A., Hundt, N. E., Cobb, A. R., Nelson-Gray, R. O., & Looten, C, M. (2007). An Analysis of Reinforcement Sensitivity Theory and the Five-Factor Model. *European Journal of Personality*, *21*(1), 869–887. https://doi.org/10.1002/per.644

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., & Weiskopf, N. (2007). When Fear Is Near: Threat Imminence Elicits Prefrontal Periaqueductal Gray Shifts in Humans. *Science*, *1119*(August), 1079–1083. https://doi.org/10.1126/science.1144298.When

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Schlund, M. W., … Cataldo, M. F. (2009). From Threat to Fear: The Neural Organization of Defensive Fear

Systems in Humans. *The Journal of Neuroscience*, *29*(39), 12236 LP-12243. https://doi.org/10.1523/JNEUROSCI.3261-16.2017

Möller, C., Wiklund, L., Sommer, W., Thorsell, A., & Heilig, M. (1997). Decreased experimental anxiety and voluntary ethanol consumption in rats following central but not basolateral amygdala lesions. *Brain Research*. https://doi.org/10.1016/S0006-8993(97)00308-9

Montoya, E. R., van Honk, J., Bos, P. A., & Terburg, D. (2015). Dissociated neural effects of cortisol depending on threat escapability. *Human Brain Mapping*, *36*(11), 4304–4316. https://doi.org/10.1002/hbm.22918

Myers-Schulz, B., & Koenigs, M. (2012). Functional anatomy of ventromedial prefrontal cortex: Implications for mood and anxiety disorders. *Molecular Psychiatry*, *17*(2), 132–141. https://doi.org/10.1038/mp.2011.88

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Sawamoto, N., … Shibasaki, H. (1999). Transient neural activity in the medial superior frontal gyrus and precuneus time locked with attention shift between object features. *NeuroImage*, *10*(2), 193–199. https://doi.org/10.1006/nimg.1999.0451

Naragon-Gainey, K. (2010). Meta-Analysis of the Relations of Anxiety Sensitivity to the Depressive and Anxiety Disorders. *Psychological Bulletin*, *136*(1), 128–150. https://doi.org/10.1037/a0018055

O'Neil, E. B., Newsome, R. N., Li, I. H. N., Thavabalasingam, S., Ito, R., & Lee, A. C. H. (2015). Examining the Role of the Human Hippocampus in Approach-Avoidance Decision Making Using a Novel Conflict Paradigm and Multivariate Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *35*(45), 15039–15049. https://doi.org/10.1523/JNEUROSCI.1915-15.2015

Ottenbreit, N. D., & Dobson, K. S. (2004). Avoidance and depression: the construction of the Cognitive-Behavioural Avoidance Scale. *Behaviour Research and Therapy*, *42*(3), 293–313.

Pellow, S., Chopin, P., File, S. E., & Briley, M. (1985). Validation of open : closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods*. https://doi.org/10.1016/0165-0270(85)90031-7

Perkins, A. M., Arnone, D., Smallwood, J., & Mobbs, D. (2015). Thinking too much: Self-generated thought as the engine of neuroticism. *Trends in Cognitive Sciences*, *19*(9), 492–498. https://doi.org/10.1016/j.tics.2015.07.003

Perkins, A. M., Ettinger, U., Weaver, K., Schmechtig, A., Schrantee, A., Morrison, P. D., … Corr, P. J. (2013). Advancing the defensive explanation for anxiety disorders: lorazepam effects on human defense are systematically modulated by personality and threat-type. *Translational Psychiatry*, *3*(4), e246. https://doi.org/10.1038/tp.2013.20

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learning and Memory*, *22*(9), 417–425. https://doi.org/10.1101/lm.039180.115

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Grillon, C., Gore, J. C., & Davis, M. (2001). Activation of the human amygdala to a cognitive representation of fear. *Nature Neuroscience*, *4*(437), 437–441.

Quirk, G. J., Likhtik, E., Pelletier, J. G., & Paré, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(25), 8800–8807. https://doi.org/23/25/8800 [pii]

Radke, S., Volman, I., Kokal, I., Roelofs, K., de Bruijn, E. R. A., & Toni, I. (2017). Oxytocin reduces amygdala responses during threat approach. *Psychoneuroendocrinology*, *79*, 160–166. https://doi.org/10.1016/j.psyneuen.2017.02.028

Radke, S., Volman, I., Mehta, P., Van Son, V., Enter, D., Sanfey, A., … Roelofs, K. (2015). Testosterone biases the amygdala toward social threat approach. *Science Advances*, *1*(5). https://doi.org/10.1126/sciadv.1400074

Radua, J., & Mataix-Cols, D. (2012). Meta-analytic methods for neuroimaging data explained. *Biology of Mood & Anxiety Disorders*, *2*, 6. https://doi.org/10.1186/2045-5380-2-6

Radua, J., Mataix-Cols, D., Phillips, M. L., El-Hage, W., Kronhaus, D. M., Cardoner, N., & Surguladze, S. (2012). A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *27*(8), 605–611.

https://doi.org/10.1016/j.eurpsy.2011.04.001

Radua, J., van den Heuvel, O. A., & Surguladze, S. (2010). Meta-analytical Comparison of Voxel-Based Morphometry Studies in Obsessive-Compulsive Disorder vs Other Anxiety Disorders. *Arch Gen Psychiatry*, *67*(7), 701–711.

Rigoli, F., Ewbank, M., Dalgleish, T., & Calder, A. (2016). Threat visibility modulates the defensive brain circuit underlying fear and anxiety. *Neuroscience Letters*, *612*, 7–13. https://doi.org/10.1016/j.neulet.2015.11.026

Rigoli, F., Pavone, E. F., & Pezzulo, G. (2012). Aversive Pavlovian Responses Affect Human Instrumental Motor Performance. *Frontiers in Neuroscience*, *6*. https://doi.org/10.3389/fnins.2012.00134

Roberts, K. L., & Hall, D. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory …. *Journal of Cognitive Neuroscience*, *20*(6), 1063–1078. https://doi.org/10.1162/jocn.2008.20074

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M. D., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage*, *136*, 94–105. https://doi.org/10.1016/j.neuroimage.2016.04.070

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., … Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience*, *27*(9), 2349–2356. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Servatius, R. J. (2016). Editorial: Avoidance: From Basic Science to Psychopathology. *Frontiers in Behavioral Neuroscience*, *10*(February), 1–4. https://doi.org/10.3389/fnbeh.2016.00015

Stein, M. B., & Paulus, M. P. (2009). Imbalance of Approach and Avoidance : The Yin and Yang of Anxiety Disorders. *Biological Psychiatry*, *66*(12), 1072–1074. https://doi.org/10.1016/j.biopsych.2009.09.023

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How Humans Integrate the Prospects of Pain and Reward during Choice. *Journal of Neuroscience*,

29(46), 14617–14626. https://doi.org/10.1523/JNEUROSCI.2026-09.2009

Terburg, D., Morgan, B. E., Montoya, E. R., Hooge, I. T., Thornton, H. B., Hariri, A. R., … Van Honk, J. (2012). Hypervigilance for fear after basolateral amygdala damage in humans. *Translational Psychiatry*, *2*(5), e115-10. https://doi.org/10.1038/tp.2012.46

Vogel, J. R., Beer, B., & Clody, D. E. (1971). A simple and reliable conflict procedure for testing anti-anxiety agents. *Psychopharmacologia*. https://doi.org/10.1007/BF00403989

Walker, D. L., & Davis, M. (2002). The role of amygdala glutamate receptors in fear learning, fear-potentiated startle, and extinction. *Pharmacology Biochemistry and Behavior*, *71*(3), 379–392. https://doi.org/10.1016/S0091-3057(01)00698-0

Warden, S. J., & Mantila Roosa, S. M. (2014). Physical activity completed when young has residual bone benefits at 94 years of age: A within-subject controlled case study. *Journal of Musculoskeletal Neuronal Interactions*, *14*(2), 239–243. https://doi.org/10.1038/nature13314.A

Wendt, J., Löw, A., Weymar, M., Lotze, M., & Hamm, A. O. (2017). Active avoidance and attentive freezing in the face of approaching threat. *NeuroImage*, *158*(June), 196–204. https://doi.org/10.1016/j.neuroimage.2017.06.054

Williams, G. V., Rolls, E. T., Leonard, C. M., & Stern, C. (1993). Neuronal responses in the ventral striatum of the behaving macaque. *Behavioural Brain Research*, *55*(2), 243–252. https://doi.org/10.1016/0166-4328(93)90120-F

Zhao, X. H., Wang, P. J., Li, C. B., Hu, Z. H., Xi, Q., Wu, W. Y., & Tang, X. W. (2007). Altered default mode network activity in patient with anxiety disorders: An fMRI study. *European Journal of Radiology*, *63*(3), 373–378. https://doi.org/10.1016/j.ejrad.2007.02.006

Table 1

Simple avoidance paradigms

Study	N (male)	Stimuli (threat)	Brief task description
Avoidance of an active predator threat through navigation of avatar			
Mobbs et al. (2007)	14	Electric shock	Maze navigation to escape virtual predator, stratified as high (3 shocks), low (1 shock) or neutral (0 shocks) threat in active condition. Predator mimicked participant avatar movement only for control trials.
Mobbs et al. (2009)	24 (12)	Electric shock	As in Mobbs et al., 2009 with two additions: visual cues of probability of capture set at 87.5% (high probability) or 12.5% (low probability) and a maze exploration incentive added (participants instructed collect yellow triangles scattered throughout maze).
Collins et al. (2014)	28 (14)	Electric shock	Participants must make navigational movements (specifically, crossing a particular part of the on screen grid) to avoid aversive outcome when threat symbol is displayed. Incorrect movements would result in aversive event. Participants were not explicitly told the correct navigational movements, but learned through trial-and-error. Motor control trials were included in which threat was absent.
Rigioli et al. (2016)	22 (11)	Loud aversive noise burst	Navigation of avatar along a pathway (towards the right) to escape predator appearing far left. Probability of capture set at 50% of trials. Trials either had a visible or an invisible predator (the latter requiring participants to escape without knowledge of predator proximity/speed).
Boeke et al. (2017)	56 (0)	Electric shock	The relationship between face stimuli and shock was taught through fear conditioning in an acquisition phase; faces were presented in pairs so it was not clear which was the threat. Following this, the faces were again presented but participants could attempt to avoid the aversive outcome associated with them by moving a circle around a grid onscreen, though they were not told which movements would prevent outcome. Participants were either ‘masters’ (made autonomous movements) or ‘yoked’ (passively viewed the movements of their paired master, receiving whatever outcome they received.

			Avoidance of threat is achieved through option selection
Montoya et al. (2015)*	18 (18)	Aversive image/unpleasant noise	Button press to avoid a threat image (sound pictogram), which rapidly grows to full-size to indicate threat approach. Threats are manipulated to be escapable, imminent (chance-level of escape) or inescapable. Aversive noise was presented as 'predator attack' and occurred in the escapable and imminent conditions if the button was not pressed in time. Inescapable trials presented aversive noise and full-size image immediately. Control condition involved the sound pictogram image with a cross through it.
Schlund et al. (2016)	30 (16)	Loss of money	Choice of two options to avoid threat, selection of incorrect option resulted in aversive outcome. Threat stratified as avoidable, unavoidable or safe (control). Participants were not explicitly told the correct choice, but learned through practice session trial-and-error.
Wendt et al. (2017)	24 (12)	Electric shock	Experiment had active (threat avoidable with fast button press) and passive (not avoidable, event occurs 50% of the time) trials. The stimuli signalling active vs. passive would grow in size after the participant made their response, indicating how close the threat was getting, culminating in aversive event (if active trial but did not press, or a passive trial with threat) or no event (active trial and pressed, or passive trial with no threat).

* Study administered cortisol or placebo to participants.

Table 2

Goal-conflict paradigms

Study	N (male)	Stimuli (threat/reward)	Brief task description
			Approach and avoidance through navigation of an avatar.
Bach et al, 2014	19 (19)	Token loss/gain	Participants collect tokens scattered around a virtual space using an avatar, whilst under threat of a predator waking up and chasing them, resulting in the loss of collected tokens. Threat level (risk of predator waking up) were stratified at 20%, 50% or 80% (these probabilities were visually signalled but not explicitly told to participants; one version of the experiment reported instead varied predator speed). Safe spaces were available, where participants could avoid the predator entirely (but not gather tokens); starting position and trial duration were varied randomly.
Aupperle et al, 2015	15 (8)	Aversive images/token gain	Navigation along a runway to indicate choice between two pictures representing outcomes (one at each end). Each outcome was an image-sound pairing - either positive (e.g. a sunshine) or negative (e.g. a cloud) image - combined with a certain level of tokens (0, 2, 4 or 6). If the participant moved to the middle of the runway, they had a 50% likelihood of each outcome; if at either extremity they had a 90% chance of nearest outcome (and 10% of furthest) etc. (so there was never certainty). Conflict trials offered 2, 4 or 6 points for approaching the negative stimuli pairing. Control trials involved simple avoidance (no points, just avoid negative stimuli pairing) and simple approach (few points offered, positive stimuli pairings at both ends).
Gonen et al, 2015	46 (24)	Token loss/gain	Participants earned tokens by catching coins and avoiding balls that interspersed them. Trials were either controlled (where participant actively approach/avoided coins/balls) or uncontrolled (where the participant was hit at random by coins and balls). Game difficulty was modified dynamically as the trials progressed. Trials were also separated in to high and low goal-conflict versions, by manipulating the number of ball the participant must avoid to get to the coins. The authors designed a slight bias towards controlled reward to ensure motivation was maintained.

Schlund et al, 2016	30 (16)	Token loss/gain	An initial acquisition phase required participants to pair increasing levels on a vertical bar with increasing probability of the stimuli occurring (probability of loss). In the main task, a reward and a threat level were presented to the participant, and they were given a choice between approaching (causing either gain or loss) or avoiding (avoiding loss, but also preventing gain) by pressing different buttons to indicate their selection.
Khemka et al, 2017[i]	25 (11)	Token loss/gain	As in Bach et al, 2017.
		Approach and avoidance through option selection	
Talmi et al, 2009	18 (6)	Electric shock/token gain	Participants were presented with a face pairing and a monetary amount. One of the faces was associated with a 75% probability of receiving an outcome (and 25% of getting nothing), and the other with 25% chance (and 75% of getting nothing). Whichever outcome resulted, participants would simultaneously receive an either electric shock or a 'touch' (a non-painful shock). A positive token amount, then, would cause conflict between token gain and electric shock avoidance. Participants started with £20, and selected which of the two faces they wanted to 'play' with.
Cunningham et al, 2011	18 (8)	Aversive/pleasant images	Participants were presented with a series positive, negative and neutral images (Lang et al, 2005), one at a time. Participants pressed one button to 'approach' and another to 'avoid' each image. Avoiding would cause the image to shrink and approaching would cause it to grow to fill the screen. Participants were told to make only one type of response in each block (i.e. approaching all images in the block, regardless of emotional valance) and then switch to a different response for the next block. This was intended to ensure equal approach and avoidance behaviour (and would also cause conflict, as negative images must be 'approached').
O'Neil et al, 2015	18 (9)	Token loss/gain[†]	In the learning phase, facial and scenery images were presented as pairs to participants, and associated as a pair with either reward or punishment (token loss/gain). Participants then saw the pairs recombined as either no-conflict positive (both originally in reward pairs), no-conflict negative (both originally punishment) or conflict (one reward image, one punishment). They had to decide whether to approach or avoid using a button press. No feedback was provided as to the outcome of their decision.
Loh et al, 2016[ii]	20 (9)	Token loss/gain	Participants were shown an onscreen grid, hiding both rewards (tokens) and threats ('bombs'). Participants could not see which were tokens, and were offered a series of choices: they could accept the grid and risk of threat potentially uncovering a reward, they could choose to 'explore' (as the cost of a number of tokens) and

			reveal what is under a portion of the grid (but not the whole grid, meaning threats may still be present) before deciding to take the risk, or they could decline the risk (thus avoiding threat, but also reward). If they accept a grid with only reward they received tokens, but if they accepted a grid with hidden threats they would lose tokens.
Radke et al, 2015*	54 (0)	Aversive/pleasant images	Participants were required to use a joystick to 'approach' (pulling towards themselves) or 'avoid' (pushing away from themselves) emotional face images presented on screen. Participants were told at the start of each block which movement they should make (approaching or avoiding); image size did not change dependent on response, unlike in other uses of this method.
Radke et al, 2017**	57 (57)	Aversive/pleasant images	As in Radke et al, 2015.

* Study administered testosterone or placebo to participants.

** Study administered oxytocin or placebo to participants.

[i] Magneticoencephalography study

[ii] Analysis did not use conventional MNI space

[†] Token loss/gain only in learning phase, not in decision making phase (only latter took place in fMRI scanner)

Table 3

Neural activation reported in threat avoidance studies

Study	Behaviour/Interaction	Neural activation	Self-report and physiological measures
		Avoidance of an active predator threat through navigation of avatar	
Mobbs et al, 2007	Defensive distance	Proximal threat: increased PAG, right dorsal amygdala (CeA, BNST; high threat trials only), dorsal anterior CC, premotor, pons Distal threat: increased vmPFC, subgenual anterior CC, right lateral amygdala (BLA; low threat trials only), medial OFC, lateral PFC, dorsal striatum*	Post-scan (high) dread was associated with enhanced PAG activity (peaking in DRN) in high and low threat. Post-scan (low) confidence in escape correlated with PAG, and high with ventromedial PFC.
	Threat level	High: Increased PAG, dorsolateral PFC*, hippocampus, CeA Low: Increased ventromedial PFC, dorsomedial PFC*, dorsolateral PFC*, BLA, fusiform gyrus	No correlation between these measures observed
	Threat presence	Increased cerebellum*, PAG* & posterior thalamus*; decreased medial PFC*, right ventromedial PFC* & amygdala*	
	Anticipation of threat	Increased right anterior CC, right medial OFC, ventromedial PFC, Premotor*	
Mobbs et al, 2009	Defensive distance	Proximal threat: increased midbrain*, mediodorsal thalamus, right striatum, right insula*, dorsal anterior CC*, parietal cortex*, cerebellum*, dorsolateral PFC	SCR was shown to increase from post-encounter to circa-strike, largest for high threat.

		Distal threat: increased posterior CC, bilateral hippocampus, hypothalamus, amygdala, ventromedial PFC, subgenual anterior CC	Anxiety measured pre-encounter, post-encounter and immediately before attack (circa-strike): threat condition was associated with higher anxiety (highest for high threat)
	Threat level	High threat: increased ventromedial PFC (pregenual ACC*), dorsomedial PFC, parietal lobule, parahippocampal gyrus	Trait anxiety (STAI) was associated with increased bilateral amygdala and anterior CC in high > low threat circa-strike
		Low threat: no significant voxels	
	Errors under threat	Heightened errors correlated with left PAG, dorsal anterior CC, right insula, right midbrain	
		Decreased errors correlated with ventromedial PFC, pregenual anterior CC, temporal pole	
Collins et al, 2014	Threat level	Threat > non-threat: increased bilateral anterior insula*, right dorsolateral PFC*, right caudate*, right PMC*	Anxiety VAS, STAI, COPE, BIS and IMT pre-scan; no correlations between any measure and number of aversive stimuli executed. Anxiety VAS was significantly higher during task (vs. pre- and post- task)
		Non-threat > threat: increased left medial PFC*, bilateral posterior insula*, left posterior CC*, left IPL*, left dorsomedial PFC*, left parahippocampal gyrus*, left SMA*, right amygdala*	
	Threat avoidance ability	Correlation between avoidance performance and: right amygdala – medial PFC connectivity, amygdala - post-central gyrus connectivity	
		Avoidance of threat through option selection	
Montoya et al, 2015	Anticipation	Threat: increased anterior insula*, midbrain*, dorsal anterior CC, supplementary motor area*, left putamen*	Fear of threat VAS and POMS recorded pre- and post- cortisol administration to assess parity of groups (no significant difference)
		Non-threat: increased medial TC*, medial OFC*	
	Threat presence	Threat > non-threat, escapable: no significant voxels; inescapable: anterior insula*, supplementary motor area*, right midbrain, dorsal anterior CC, medial PFC, decreased PFC, medial TC.	Speed of escape as a measure of threat sensitivity, higher in threat relative to safe conditions

	Defensive distance	Threat > non-threat, imminence: anterior insula, midbrain	
Wendt et al, 2017	Defensive distance & threat level	Proximal (high > low threat) increased bilateral anterior insula, ventrolateral PAG, dorsolateral PAG; decreased bilateral amygdala, ventromedial PFC Distal (low > high threat): increased anterior insula, dorsolateral PAG, ventromedial PFC	Heart rate, electrodermal activity and startle blink reflex assessed during training; electrodermal activity higher in anticipation of threat, increasing linearly with threat imminence. Startle potentiation was associated with threat relative to safe conditions
	Receiving aversive stimuli	Increased anterior, middle & posterior insula, anterior & middle cingulate gyrus	
Rigoli et al, 2016	Threat visibility	Hidden threat: increased bilateral hippocampus, bilateral amygdala, ventromedial PFC, posterior cingulate*, cuneus*, lingual gyrus* Visible threat: increased PAG, left inferior temporal gyrus*, precuneus*, medial frontal gyrus*, fusiform gyrus*, inferior frontal gyrus*, left cerebellum*	Non-significant correlation between trait anxiety (STAI) and probability of escape VAS rating. Positive association between STAI score and left hippocampal activation in hidden threat condition; positive association between probability of escape VAS and left hippocampal at trial start, but not end
Schlund et al, 2016	Threat level	Unavoidable threat > non-threat: increased dorsal anterior CC, dorsomedial PFC Unavoidable threat > avoidable: increased left dorsomedial PFC, left pregenual anterior CC, right ventromedial PFC	Threat ratings higher after threat conditioning; ratings for avoidable threat were significantly lower than for unavoidable threat, but safe condition significantly lower than both
Boeke et al, 2017	Threat level	Threat vs. no threat: increased putamen, caudate, medial PFC[†]	SCR was not shown to correlate with behaviour or with feelings of control during task. STAI, IUS and PSS measured prior to task, to assess parity between groups

PAG, Periaquaductal Grey; CeA, Central Amygdala; BNST, Basal Nuclei of Stria Terminalis; CC, Cingulate Cortex; PFC, Prefrontal Cortex; BLA, Basolateral Amygdala; OFC, Orbitofrontal Cortex; DRN, Dorsal Raphe Nucleus; SCR, Skin Conductance Response; STAI, State-Trait Anxiety Inventory; TC, Temporal Cortex; VAS, Visual Analogue Scale; POMS, Profile of Mood States questionnaire; SMA, Supplementary Motor Area; BIS, Barrat's Impulsivity Scale; IMT, Intrinsic Motivation Inventory; IUS, Intolerance of Uncertainty; PSS, Perceived Social Stress scale.

*Significant at <.05 corrected for multiple comparisons (i.e. surviving whole-brain analysis correction)

† Paradigm compared master (active) and yoked (passive) participants, interactions reported are master > yoked; yoked > masters revealed vmPFC activation increase (ROI analysis only).

Table 4
Neural activation in goal-conflict studies

Study	Behaviour/Interaction	Neural activation	Self-report associations
		Approach and avoidance through navigation of an avatar	
Aupperle et al, 2015	Conflict vs. no conflict	Conflict: increased right rostral & dorsal CC, right dorsolateral PFC, bilateral interior insula & bilateral caudate No conflict: increased bilateral posterior insula, dorsal mid-cingulate, left lateral PFC	No significant association between trait anxiety (STAI) and task performance
Schlund et al, 2016a	Conflict level	Smaller difference in outcome value: increased pregenual & dorsal anterior CC, dorsal & ventral cingulate, anterior insula and inferior frontal regions Larger difference in outcome value: increased ventromedial PFC and dorsolateral PFC At threshold between outcome values: peak activation, OFC and ventral hippocampus	SCR to threat stimuli increased after conditioning (in a separate experiment to neuroimaging, within same publication)
Bach et al, 2014	Threat level	As threat level increased, increased activity in the hippocampus, extending in to posterior amygdala, left parahippocampal gyrus*, left fusiform/parahippcampal gyrus*, right inferior frontal gyrus/insula*	N/A
Gonen et al, 2015	Conflict level[ι]	High conflict: increased left VTA*, bilateral pulvinar*, bilateral precuneus*, bilateral occipital lobe*, bilateral	SCR significantly higher in active threat trials.

		premotor cortex*, bilateral VS*, right MGF*, PAG* Low conflict: increased bilateral STG*, bilateral ventrolateral PFC*, right SFG*, right IFC*, medial PFC*, posterior CC*	Grouped participants as high approach or high avoidance, depending on NEO-FFI, TPQ and SPSRQ scores; significant difference between groups in VTA and VS activation, with higher activation in approach group during high goal conflict.
Khemka et al, 2017[2]	Threat level	High threat: increased right hippocampal oscillation, decreased left hippocampal oscillation Token appearance (high > low threat): increased bilateral MFG oscillation	N/A
		Approach and avoidance through option selection	
Talmi et at, 2009	Sensitivity to pain	High pain: increased activity in somatosensory cortex*	SCRs in response to pain were higher in trials with zero or negative reward, indicating pain attenuated response to reward
	Reward prediction[3]	Errors: increased activity in ventromedial PFC, OFC, anterior & posterior CC, VS, insula, hippocampus/amygdala, SFG/MFG, fusiform gyrus, rolandic operculum	
Cunningham et al, 2014	Stimuli valence	Increased right amygdala activity for positive and negative stimuli, relative to neutral	Grouped participants by BFAS score in to Neuroticism-withdrawal or Neuroticism-volatile; former predicted amygdala response to approached stimuli, latter associated with amygdala response to negative stimuli Measured BIS/BAS self-report, but no associations observed
	Motivational direction	Approach > avoid: increased amygdala activity	
O'Neil et al,	Conflict vs. no conflict	Conflict: increased bilateral posterior hippocampus*,	

2015		posterior cingulate gyrus*, paracingulate gyrus*, frontal pole*, OFC*, anterior CC*, amygdala*, putamen*, caudate*	N/A
	Motivational direction	Approach: increased anterior hippocampus*, parahippocampal cortex*, paracingulate gyrus*, temporal fusiform cortex*, OFC, insular cortex*, thalamus*, frontal pole*, inferior temporal gyrus* & entorhinal cortex*	
Radke et al, 2015	Conflict vs. no conflict	Conflict: increased anterior PFC activity (testosterone & placebo)	
	Motivational direction	Testosterone in approach: increased right amygdala activity Testosterone in avoid: decreased amygdala activity	Recorded NEO-FFI, BIS/BAS, STAI, IRI and PRF means, but no interactions explored
Radke et al, 2017	Conflict vs. no conflict	Conflict: increased anterior PFC* activity, left MTG*, left medial FC*, left inferior parietal lobule*, left paracentral lobule*, bilateral postcentral gyrus*, right STG* (all oxytocin & placebo)	LAS measure of anxiety recorded post-scan, but no significant associations shown
	Motivational direction	Approach: decreased right amygdala (oxytocin); no effect in avoidance	
Loh et al, 2017	Decision making	Avoidance of loss: bilateral inferior hippocampus activation Choice phase: increased bilateral dorsolateral PFC*, parietal cortex*, cerebellum*, right striatum*, occipital cortex*, insula*, bilateral hippocampus* Exploration phase: increased right striatum*, rostrolateral frontopolar cortex*, MFG *, SFG* & paretial cortex*	Trait anxiety (STAI) shown to correlate with behaviour in response to conflict (i.e. likelihood of choosing to avoid threat in approach-approach and of accepting offer in approach-avoidance

VTA, Ventral Tegmental Area; VS, Ventral Striatum; CC, Cingulate Cortex; PFC, Prefrontal Cortex; OFC, Orbitofrontal Cortex; MFG, Middle Frontal Gyrus; PAG, Periaquaductal Grey; STG, Superior Temporal Gyrus; SFG, Superior Frontal Gyrus; IFC, Inferior Frontal Cortex;, STAI, State-Trait Anxiety Inventory; SCR, Skin Conductance Response; NEO-FFI, NEO Five Factor Inventory; TPQ, Tri-dimensional Personality Questionnaire; SPSRQ, Sensitivity to Punishment and Reward Questionnaire; BFAS, Big Five Aspect Scale; BIS/BAS, Behavioural Inhibition System/Behavioural Activation System Scales; IRI, Interpersonal Reactivity Index; PRF, Personality Research Form.

[1] Only approach behaviour in controlled trials analysed due to lack of avoidance behaviours

[2] Magentoencephalography (MEG) study

[3] Analysis restricted to areas activated by main effect of pain (whole-brain) across the whole group

Table 5

Details of studies included in SDM pooled data analysis

Study	N	Contrasts	Thresholding	Data type
Mobbs et al, 2009	24	Escape from predator, high vs. low threat*	<.001 uncorrected	Whole-brain coordinates
Bach et al, 2014†	19	Escape from predator, high vs. low threat*	<.05 corrected	SPM t-map
Collins et al, 2014	28	Escape from threat, threat vs. non-threat*	<.05 corrected	Whole-brain coordinates
Wendt et al, 2017	24	Avoiding threat, threat vs. non-threat**	<.001 corrected	SPM t-map
Boeke et al, 2017††	56	Avoiding threat, threat vs. non-threat**	<.05 corrected	SPM t-map

SDM, Seed-based d Mapping; SPM, Statistical Parametric Mapping

† Goal-conflict paradigm, but simple avoidance available for SDM analysis
†† Study used FSL stereotactical space
* Movement of on screen avatar to escape threat
** Button press to signal decision to avoid threat

Table 6

Areas of activation and deactivation identified in SDM pooled-data analysis, high threat vs. low- or absent-threat.

Region	Peak MNI coordinate	SDM Z-value	*P*	Voxels	BA	Heterogeneity[7] (n = 5 studies)
Activations						
R superior frontal gyrus, medial[1]	4, 38, 38	9.094	<.001	2664	9	0
R superior frontal gyrus, dorsolateral	22, 42, 38	3.960	<.001	35	9	0
R superior frontal gyrus, dorsolateral	20, 56, 24	3.065	.002	11	10	5.2
R middle frontal gyrus[2]	40, 28, 36	4.317	<.001	1735	45	0
R middle frontal gyrus[1]	28, 52, 22	3.564	<.001	141	46	0
L superior longitudinal fasciculus III	-38, 16, 10	4.063	<.001	705	-	0
R superior longitudinal fasciculus III	44, -48, 36	3.276	.001	19		5.8
R supramarginal gyrus[3]	56, -44, 32	3.628	<.001	247	48	4.3
Cerebellum, vermic lobule IV/V	0, -64, -6	3.052	.002	29		0
Deactivations						
L superior temporal gyrus[4]	-60, -12, 10	-3.832	<.001	1184	22	12.6
R fronto-insular tract 5	52, -10, 18	-3.556	<.001	821	-	0
L gyrus rectus	-4, 46, -20	-3.261	<.001	854	11	3.5

L precuneus	-10, -54, 30	-3.259	<.001	626	23	20.9
L middle frontal gyrus[5]	-22, 26, 44	-3.420	<.001	303	9	0
L superior frontal gyrus, dorsolateral	-14, 62, 16	-2.398	<.001	45	10	0
L inferior frontal gyrus, triangular part	-50, 28, 20	-2.33	<.001	37	45	0
L middle frontal gyrus, orbital part	-22, 36, -16	-2.472	<.001	24	11	0.4
L middle temporal gyrus	-60, -6, -10	-2.312	<.001	41	22	0.6
L middle temporal gyrus	-52, -32, 8	-2.406	<.001	11	22	0
L inferior temporal gyrus	-56, -48, -12	-2.476	<.001	20	20	62.3
R temporal pole, superior temporal gyrus[6]	40, 6, -22	-2.712	<.001	31	38	12.2
R cerebellum, crus I	28, -72, -36	-2.347	<.001	25	-	0
R cerebellum, crus II	46, -68, -40	-2.130	.002	16	-	0
R parahippocampal gyrus	16, -6, -26	-2.521	<.001	7	28	0
Olfactory cortex	2, 12, -10	-2.467	<.001	3	25	0

R, Right; L, Left; MNI, Montreal Neurological Institute; SDM, Seed-based d Mapping; BA, Brodmann Area

[1] Driven by 3/5 studies: Boeke et al, 2017, Collins et al, 2014, Bach et al, 2014

[2] Driven by 4/5 studies: Boeke et al, 2017, Bach et al, 2014, Wendt et al, 2017, Collins et al, 2014

[3] Driven by 4/5 studies: Bach et al, 2014, Wendt et al, 2017, Mobbs et al, 2009, Collins et al, 2014

[4] Driven by 4/5 studies: Bach et al, 2014, Boeke et al, 2017, Mobbs et al, 2009, Collins et al, 2014

[5] Driven by 4/5 studies: Bach et al, 2014, Boeke et al, 2017, Wendt et al, 2017, Mobbs et al, 2009

[6] Driven by 4/5 studies: Mobbs et al, 2009, Wendt et al, 2017, Collins et al, 2014

[7] Heterogeneity scores (0-100) calculated in MRIcron.

Brain activation and threat